

DC5 Pema FRICK

IRP5: Multilingual gender fairness in AI-generated texts

Research fields: Applied Linguistics, Gender Linguistics, Contrastive Linguistics, Generative Artificial Intelligence (GenAI)

Co-supervisors of IRP5:

Mirjam Schmuck, mirjam.schmuck@hum.ku.dk

Hélène Vinckel-Roisin, helene.vinckel-roisin@univ-lorraine.fr

This PhD position is part of a double-degree programme. Upon completion, candidates will be awarded two PhD diplomas, following research stays at both institutions: **University of Copenhagen** (diploma conferred by the **University of Copenhagen**) and **Université de Lorraine** (diploma conferred by the **Université de Lorraine**).

Scientific context of the Individual Research Project IRP5:

IRP5 is embedded into **Work Package 1 (WP1)** ‘Multilingual communication and resources in the Digital Age: discourses, practices, stereotypes’. WP1 investigates how the concepts of Language Awareness and pluri-/multilingualism are discursively constructed, ideologically framed, and technologically mediated in digital contexts across various societal domains (e.g., media, education, industry, politics). It will provide a research-based catalogue of arguments and language ideologies, with particular attention to issues of linguistic justice, gender fairness, and inclusion.

In this context, IRP5 investigates multilingual gender fairness in AI-generated texts.

Recruiting University: University of Copenhagen (UCPH), Denmark

Partner University for the Joint Doctorate: Université de Lorraine (UL), Nancy, France

Objectives:

Despite major advances in GenAI, stereotyping and biases are ubiquitous in texts generated by AI technologies, which runs counter to a European society committed to gender equality. Working in close collaboration with DC4, DC5 will:

- 1) investigate texts generated by leading LLMs (e.g. OpenAI’s ChatGPT, Google’s Gemini, Mistral’s Le Chat) in terms of gender fairness with a focus on languages other than EN (e.g. DA, FR, DE);
- 2) explore the impact of linguistic factors (i.e. linguistic bias), specifically the effects of grammatically un/gendered languages and the masculine default, on gendered representations using corpus linguistic methods;
- 3) reveal how AI tools echo and reinforce human bias as well as raise Critical Multilingual Language Awareness among users and developers of how LLMs index gender-related structures of power;
- 4) contribute to enhancing the dialogue between linguists and computer scientists to confront sexism in GenAI technologies with the Associated Partners.

Expected Results:

- Dataset of AI-generated texts in three languages;
- In-depth analysis of gendered representations in AI-generated texts and the impact of linguistic factors;
- Proposal of language-specific measures to confront algorithmic discrimination against women in AI technologies to reduce harm caused by biased systems;
- Two podcasts aimed at non-expert audience during the secondment at RHET-AI.

Secondments:

- Center for Rhetorical Science Communication Research on Artificial Intelligence, Unit 1, Karlsruhe (Germany), 4 months
 - Leibniz Institute for the German Language (IDS), Mannheim (Germany), 1 month

Languages: English, German, French.